

University of Groningen

## Value of Functional Capacity Evaluation Information in a Clinical Setting for Predicting a Return to Work

Streibelt, Marco; Blume, Carsten; Thren, Karsten; Reneman, Michiel F.; Mueller-Fahrnow, Werner

*Published in:*  
Archives of Physical Medicine and Rehabilitation

*DOI:*  
[10.1016/j.apmr.2008.08.218](https://doi.org/10.1016/j.apmr.2008.08.218)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2009

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Streibelt, M., Blume, C., Thren, K., Reneman, M. F., & Mueller-Fahrnow, W. (2009). Value of Functional Capacity Evaluation Information in a Clinical Setting for Predicting a Return to Work. *Archives of Physical Medicine and Rehabilitation*, 90(3), 429-434. <https://doi.org/10.1016/j.apmr.2008.08.218>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

## ORIGINAL ARTICLE

# Value of Functional Capacity Evaluation Information in a Clinical Setting for Predicting Return to Work

Marco Streibelt, PhD, Carsten Blume, Dip, Karsten Thren, PhD, Michiel F. Reneman, PhD, Werner Mueller-Fahrnow, MD, PhD†

**ABSTRACT.** Streibelt M, Blume C, Thren K, Reneman MF, Mueller-Fahrnow W. Value of functional capacity evaluation information in a clinical setting for predicting return to work. *Arch Phys Med Rehabil* 2009;90:429-34.

**Objective:** To evaluate the quality of Functional Capacity Evaluation (FCE) information in predicting return to work (RTW).

**Design:** Prospective cohort study.

**Setting:** Inpatient rehabilitation clinic.

**Participants:** Patients (N=220) with chronic musculoskeletal disorders (MSD) conducting a medical rehabilitation.

**Interventions:** Not applicable.

**Main Outcome Measures:** Patients filled in questionnaires at admission and 1-year follow-up. An FCE was performed on admission. RTW was defined as a combination of employment at 1-year follow-up with a maximum of 6 weeks sick leave because of MSD in the postrehabilitation year. As predictive FCE information, the physical capacity (Dictionary of Occupational Titles categories 1–5), the number of test results not meeting work demands (0–25), and the tester's recommendation of work ability in the actual job ( $\geq 6$ h/d) were analyzed. Logistic regression models (crude and adjusted for the concurrent predictors employment, preadmission sick leave, and patient's prognosis of RTW) were created to predict RTW.

**Results:** Complete data were obtained for 145 patients. The sample showed a non-RTW at 1-year follow-up for 37.9%. All FCE information showed significant relations to RTW ( $r=.28-.43$ ;  $P<.05$ ). In the crude as well as in the adjusted regression models, all FCE information predicted RTW, but the models' quality was low. The integration of FCE information led to an increase of 5%. The predictive efficiency was poor. The adjusted model for failed tests showed a substantial improvement compared with the reference model (concurrent predictors only).

**Conclusions:** There was a significant relation between FCE information and RTW with and without concurrent predictors, but the predictive efficiency is poor. Primarily, the number of failed tests seemed to be of significance for patients with ambiguous RTW prognosis. A first proposal for a prediction rule was discussed.

**Key Words:** Musculoskeletal diseases; Rehabilitation.

© 2009 by the American Congress of Rehabilitation Medicine

**P**ATIENTS WITH MUSCULOSKELETAL disorders lasting longer than 3 months have an increased risk of long-term sick leave and permanent work-related disability. This represents a substantial economic problem.<sup>1-3</sup> In this context, an RTW depends on diverse personal and environmental factors. Particularly imminent or already existing chronic pain considerably increases the risk to remain off work.<sup>4-7</sup>

Appropriate problem-oriented rehabilitation can significantly reduce long-term sick leave.<sup>8,9</sup> However, the medical diagnosis is often not sufficient to identify possible future problems concerning occupational participation or to develop the necessary forms of therapy. It is important to identify relevant risk factors and deficits at an early stage because they indicate that a patient might not be able to return to work because of the MSD. The identification of persons affected by unsuccessful RTW could theoretically be facilitated by FCEs. FCEs are defined as batteries of standardized tests designed to assess systematically a person's work-related functional capacity.<sup>10,11</sup> Their development goes back to the 1970s.<sup>12,13</sup> FCEs may be applied in workers' compensation claims. Thus, the use of FCEs should provide the basis for a realistic evaluation of a person's capacity to work and of future employment opportunities. A detailed evaluation of functional capacity and deficits can also be applied in rehabilitation—for example, in planning and monitoring therapeutic interventions.<sup>14</sup>

The IWS FCE—currently known as WorkWell Systems FCE—was applied in this study. About 75% of all rehabilitation facilities in Germany applying FCE use the IWS FCE.<sup>14</sup> The test-retest reliability<sup>15-17</sup> and the interrater reliability<sup>18</sup> of FCEs were estimated as good or very good. Acceptable results concerning construct validity compared with self-assessed functional capacity were found.<sup>19-21</sup>

The application of IWS FCE in a clinical setting is effective in the assessment of individual activity limitations and consequently the therapeutic measures. A randomized controlled trial<sup>22</sup> showed that patients performing function-centered multidisciplinary rehabilitation, based on FCE results, had better results than multidisciplinary rehabilitation only. Within the

From the Department of Health Care Research and Quality Management in Rehabilitation, Charité University Berlin, Germany (Streibelt, Blume, Mueller-Fahrnow); Klinik Niedersachsen Bad Nenndorf, Germany (Thren); Center for Rehabilitation, University Medical Center Groningen, University of Groningen, The Netherlands (Reneman).

Supported by the Erwin-Roever-Foundation.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Reprint requests to Marco Streibelt, PhD, Dept of Health Care Research and Quality Management in Rehabilitation, Charité University Berlin, Luisenstr 13a, D-10098 Berlin, Germany, e-mail: [marco.streibelt@charite.de](mailto:marco.streibelt@charite.de).

†Deceased.

0003-9993/09/9003-0051\$36.00/0

doi:10.1016/j.apmr.2008.08.218

## List of Abbreviations

DOT	Dictionary of Occupational Titles
FCE	Functional Capacity Evaluation
IWS FCE	Isernhagen Work System Functional Capacity Evaluation
MSD	musculoskeletal disorder
PRE <sub>ref</sub>	proportional reduction of error
RTW	return to work
WHO	World Health Organization

1-year follow-up, an improvement in occupational participation and pain management could be verified.

Hence, IWS FCE may be considered a valid and reliable instrument for assessing the functional capacity and for providing important information regarding the enhancement of more effective multidisciplinary rehabilitation.<sup>23</sup> However, there is only limited evidence that FCE information can predict the time until RTW.<sup>24,25</sup> Consistent effects regarding a sustained RTW (eg, no additional negative occupational occurrences during follow-up) could not be proved.<sup>26,27</sup> Other authors conclude that FCE can predict the occupational status only to a limited extent.<sup>28,29</sup> There are indicators that competitive constructs can predict the occupational status to at least the same degree.<sup>30,31</sup> Especially patients' self-reported measures have been shown to be significant predictors of successful RTW. So far, 2 studies have confronted FCE-based predictions with self-reported measures: first, the self-reported pain intensity,<sup>28</sup> and second, the patient's expected disability in the job.<sup>27</sup> In both studies, a poor predictive quality of FCE tests is shown after adjusting for self-reported measures. Therefore, it is indicated that FCE provides little additional information for predicting RTW status after rehabilitation.<sup>28</sup>

However, a number of questions remain unanswered. First, previous studies predominantly used data from legal proceedings as variables for RTW. The problem arising from the underestimation of the actual RTW quotas may have occurred more in this environment than in a nonlegal rehabilitation setting.<sup>27,32</sup> Second, to date the physical capacity (often maximum weight tested in floor-to-waist lifting) has been used as potential predictor. It may be expected, however, that further information from the FCE records will be able to predict occupational participation after rehabilitation more precisely. Beyond mere test scores, personal contextual factors (eg, pain management, ergonomic handling, and potential medical or occupational interventions after rehabilitation) may be more closely related to the complex phenomenon of occupational reintegration.

For these reasons, we aimed at studying the predictive quality of FCE information regarding future occupational participation, as well as presenting these findings in contrast with concurrent predictors.

## METHODS

### Study Design

This study was considered a prospective cohort study and part of a comprehensive study to evaluate an FCE-based multidisciplinary rehabilitation for employable persons with MSDs. It was conducted at the Klinik Niedersachsen, an inpatient rehabilitation clinic near Hanover in Northern Germany.<sup>33</sup> Between July 2002 and June 2003, all patients covered by regional German statutory pension insurance with imminent or prevailing occupational disability because of MSDs were integrated in this study. This regional pension insurance mainly deals with blue collar workers. All patients filled out a questionnaire including instruments measuring health-related constructs concerning MSDs on admission, at the end of rehabilitation, and at the 6-month and 12-month follow-ups.<sup>33</sup>

The IWS FCE was conducted on 2 consecutive days during the first rehabilitation week and guided by physiotherapists with the necessary additional training. (All physiotherapists have an IWS FCE license. In Germany this is given out by the IWS FCE Academy.) The IWS FCE reflects work-related activities such as lifting, carrying, and bending. It consists of 29 standardized tests that are measured and interpreted to obtain a patient's individual physical work-related capacity. The job

demands were identified by a guided interview. The patient's functional capacity and job demands were then compared according to IWS FCE protocols (based on the DOT<sup>34,35</sup>). The resulting FCE report contained the single test results as well as the tester's rating for the actual work ability (h/d) and prospective occupational participation. To enable a comparison between FCE and self-reported measures, the FCE information was matched with the patients' questionnaires.

### Measures

Sample characteristics were measured with commonly used instruments—for example, the Medical Outcome Study 36-Item Short Form Health Survey<sup>36</sup> measuring self-reported health status, the Pain Disability Index<sup>37</sup> measuring self-reported pain-related disability, and a Numeric Rating Scale measuring pain intensity.<sup>38</sup>

**FCE report.** Three FCE-based sources of information were included in the analysis. The maximum functional capacity was measured by using kinesiophysical FCE—for example, the tests were done with steadily increasing weights until the patient showed clear physiologic signs of personal maximum ability. FCE scores of 8 tests were transformed into DOT classification (category 1, sedentary; 2, light; 3, medium; 4, heavy; 5, very heavy).<sup>20</sup> Seven tests were already available in DOT classification. This classification provides average physical capacity for every patient (DOT scale, 1–5). The rating of physical capacity was compared with job demands. For each of the 25 tests, a rating was given for whether the specific value corresponding to job demands was met. A recommendation of work-related capacity was made based on the total of deficits (number of failed tests, 0–25).<sup>39</sup> Finally, the tester's rating of actual capacity for his/her last job was stated in the FCE report. The rating, based on test scores, provided the recommendation of work ability for less than or at least 6 hours a day.

**Concurrent predictors.** Potential concurrent predictors were used to test or exclude competitive hypotheses. Literature concerning successful RTW for patients with acute MSD provided much information.<sup>40</sup> Indeed, there were limited predictors for patients with chronic disorders. Among those, the patient's expected disability in the job was considered an effective predictor.<sup>31</sup> This construct was measured by the following question: "Do you think that your performance in the job is limited due to your health status in the long term?" (not limited, partly limited, or heavily limited). Further, 2 stable predictors closely related to samples of the German rehabilitation system were included in the analysis: employment status at admission and sick leave 1 year prior to admission in weeks.<sup>6,30</sup>

**Outcome.** The outcome of this study was a successful occupational participation 1 year after the rehabilitation process. The term *participation* was based on the WHO's biopsychosocial model of the *International Classification of Functioning, Disability and Health*.<sup>41</sup> The WHO definition of occupational participation referred to 2 main aspects: the access to the job market, and the prospect of adequate participation or RTW in good health. Accordingly, 2 conditions for RTW in good health are necessary<sup>42</sup>: employment status at the 1-year follow-up, and low levels of sick leave during follow-up. In this study, all those who were employed at 1-year follow-up and who were sick-listed for a period of 6 weeks or less because of MSD were considered to have a successful occupational participation.

### Analysis

Firstly, bivariate correlation between the outcome (RTW) and the FCE information (DOT scale, number of failed tests,

**Table 1: Patient Characteristics (n=145)**

Characteristics	Mean $\pm$ SD	Percentage
Age (y)	47.9 $\pm$ 8.5	
Sex (men)		78.6
Diagnosis		
Dorsopathies		77.3
Arthropathies		12.1
Soft tissue disorders		8.3
Other		2.3
Employment status at admission (unemployed)		8.2
Sick-listed at admission		54.4
Time of sick leave, 1 y preadmission (>3mo)		36.4
Function (MOS-36, 0–100)	55.6 $\pm$ 21.9	
Pain intensity (NRS, 0–10)	6.9 $\pm$ 1.7	
Disability (PDI, 0–70)	31.2 $\pm$ 14.5	
Patients' prognosis of expected work disability (heavily limited)		44.3
FCE protocol		
DOT scale (1–5)	4.1 $\pm$ 0.4	
Number of failed tests (0–25)	4.0 $\pm$ 5.1	
Recommendation of work ability		
0–6h/d		22.2
>6h/d		77.8

Abbreviations: MOS-36, MOS 36-Item Short Form Health Survey; NRS, Numeric Rating Scale; PDI, Pain Disability Index.

recommendation of work ability) and concurrent predictors (employment status at admission, sick leave 1 year prior to admission, patient's prognosis of expected disability in the job) were calculated. For dichotomous variables, the chi-square test was used, and for metric variables, an analysis of variance was estimated. For the quantity of the relation between outcome measures and the FCE information correlation, coefficients eta and phi were used (concurrent predictors respectively). Within the multivariate analysis, a reference model (multiple logistic regression) was specified testing the influence of the predictors on the outcome. Further models were estimated containing isolated estimations of the FCE information (see crude models), and the FCE information included concurrent predictors (see adjusted models.) The influence of corresponding FCE information on the outcome was measured with a *P* value of 5% in both the crude and adjusted models. To rate the models' quality compared with each other and the reference model, the  $R^2_L$  by Hosmer and Lemeshow<sup>43</sup> was used.  $R^2_L$  was defined by the relation of log-likelihood of the adjusted model and the basic model (only intercept). The predictive efficiency was

measured by using  $\lambda_p$  and  $PRE_{ref}$ .  $\lambda_p$  was a criterion for the efficiency of the model concerning the samples' reclassification. The number of cases of the outcomes' smaller category was compared with its sum and the models' mismatched cases. The measured value was between 0 and 1. The  $PRE_{ref}$  showed the reduction of an improved model's mismatched cases compared with a basic model. It could be interpreted as part of the improvement of the predictive efficiency compared with an estimated reference model.<sup>44</sup>

## RESULTS

Of 494 patients who were consecutively assigned in the study, a total of 220 patients completed the FCE test. Of these, data at the 1-year follow-up were available for 145 patients. Table 1 shows the characteristics of the patient sample in detail.

At the 1-year follow-up, 62.1% of the patients were successfully reintegrated into work. Thus, the percentage of patients with unsuccessful occupational participation was 37.9%. All parameters of the FCE report showed a significant relation to the outcome (table 2). The correlations varied between .28 (number of failed tests) and .43 (recommendation of work ability). The FCE report showed lower measures of physical capacity and increased work-related deficits for those not returning to work. The testers estimated that fewer patients were able to recommence full-time work (>6h/d).

Additionally, the concurrent predictors showed significant differences between successful and unsuccessful occupational participation. The main difference that could be shown was in the term of the patient's expected disability at work: 30% of the successfully reintegrated persons anticipated a heavy disability, whereas about 70% of the unsuccessfully reintegrated persons held that opinion. The least difference was seen in the employment status at admission. The number of unemployed in the group of the successfully reintegrated persons was 10% lower than in the group of unsuccessfully reintegrated persons (4% vs 14%).

The multivariate analysis confirmed the bivariate analysis (table 3). All 3 crude models showed a significant relationship between the FCE information and the occupational participation at the 1-year follow-up. A 1-point increase on the DOT scale reduced the risk of non-RTW by 84%, whereas an increase of 1 failed test improved the risk by 15%. The testers' recommendation concerning an expected work ability of more than 6 hours a day lessened the risk of non-RTW by 90%. The significant relationship remained stable when the concurrent predictors were considered (see adjusted models).

On the basis of FCE information, between 8% (number of failed tests) and 17% (recommendation of work ability) of variance in outcome could be explained. Thus the crude models were in the range of the reference model (16%), respectively.

**Table 2: Relationships of FCE-Based Indicators of Performance and Concurrent Predictors to RTW: Bivariate Analysis**

	RTW	Non-RTW	<i>r</i> *	<i>P</i> ( $\chi^2$ )
FCE protocol				
DOT scale (1–5), mean $\pm$ SD	4.2 $\pm$ 0.4	3.9 $\pm$ 0.4	0.30 <sup>†</sup>	<.001 <sup>‡</sup>
Number of failed tests (0–25), mean $\pm$ SD	2.9 $\pm$ 4.5	6.0 $\pm$ 5.8	0.28 <sup>†</sup>	.001 <sup>‡</sup>
Recommendation of work ability, >6h a day (%)	90.5	53.2	0.43	<.001
Predictors				
Patients' prognosis of expected work disability (heavily limited) (%)	30.1	67.3	0.36	<.001
Time of sick leave, 1 year preadmission (>3mo) (%)	26.3	55.3	0.29	.001
Employment status at admission (unemployed) (%)	3.7	14.6	0.20	.024

\*Correlation coefficient  $\phi$ , except <sup>†</sup> $\epsilon$ , <sup>‡</sup>F test.



Table 3: Relation of FCE-Based Indicators of Performance and Concurrent Predictors to RTW: Multiple Logistic Regression Models

Model	OR (95% CI)	−2Log-Likelihood	$R^2_{\text{L}}$ <sup>  </sup>	% Correct <sup>¶</sup>			$\lambda_{\text{p}}$ (PRE <sub>ref</sub> ) <sup>#</sup>
				All	Sensitivity	Specificity	
0. Reference model							
“confounder”* (n=118)		131.620	15.6	70.3	63.6	74.3	0.21
I. Model “DOT scale” <sup>†</sup> (n=118)							
Crude	0.14 (0.05–0.38)	138.277	11.3	66.1	56.8	71.6	0.09 (−0.14)
Adjusted	0.22 (0.07–0.67)	124.044	20.4	72.9	68.2	75.7	0.27 (0.09)
II. Model “failed tests” <sup>‡</sup> (n=115)							
Crude	1.15 (1.06–1.24)	140.286	8.3	70.4	47.7	84.5	0.23 (0.03)
Adjusted	1.10 (1.01–1.19)	124.880	18.4	77.4	70.5	81.7	0.41 (0.26)
III. Model “recommended work ability” <sup>§</sup> (n=111)							
Crude	0.10 (0.03–0.29)	123.125	14.5	75.7	43.6	93.1	0.31 (0.23)
Adjusted	0.24 (0.07–0.85)	114.407	20.5	73.9	53.8	84.7	0.26 (0.17)

NOTE. Dependent variable: RTW (0, returned to work without long sick leave in the 1-year follow-up; 1, not returned/returned, but with long sick leave in the 1-year follow-up).

Abbreviations: CI, confidence interval; OR, odds ratio; PRE<sub>ref</sub>, (cases incorrectly predicted by the reference model-cases incorrectly predicted by the model)/cases incorrectly predicted by the reference model;  $\lambda_p$ , cases in the smaller category-cases incorrectly predicted by the model/cases in the smaller category.

\*Included predictors (also for adjusted models): employment status at admission (working vs not working); sick leave, 1-year preadmission (0–3mo vs >3mo); patients' prognosis of expected work disability (not/partially vs heavily limited).

<sup>†</sup>DOT scale (1–5).

<sup>‡</sup>Number of failed tests (0–25).

<sup>§</sup>Therapists' recommendation of work ability (0, 0–6h/d; 1, >6h/d).

<sup>¶</sup>Models' quality: regression sum-of-squares to the total sum-of-squares:  $R^2_L = 100(1 - L_p/L_0)$ , where  $L_0$  is the log-likelihood for the model "only intercept" and  $L_p$  the log-likelihood for the model "intercept + predictors."

<sup>#</sup>Cut-off:  $P = .379$  for non-RTW.

<sup>#</sup>Predictive efficiency: (A)  $\lambda_p$ , (B) PRE<sub>ref</sub>.

Further improvement by integrating additional FCE information into the adjusted models was limited: the  $R^2_L$  showed approximately 5% additionally explained variance in the 3 models.

The correct reclassification by the crude models was successful for 66% (DOT scale) to 76% (recommendation of work ability). By using the information of the number of failed tests, a substantial improvement of the efficiency beyond the reference model was possible. Compared with the reference model, the corresponding adjusted model showed a reduction of misclassification by 26% and had an overall efficiency of  $\lambda_p$  equal to .41.

## DISCUSSION

Within the sample used in this study, various FCE information could be considered as valid for predicting occupational reintegration. However, the predictive ability of the models was poor. The predictive ability of directly acquired information like FCE test results in DOT categories and the job requirement qualified by test results was lower than the FCE reports-based recommendations of the prospective work ability. The latter could predict occupational participation at the 1-year follow-up just as well as other predictors.

All FCE information showed a significant relationship to the prospective RTW, even when adjusted for the mentioned predictors. However, the predictive efficiency of the models was low. Additional relevant findings expressed in increased predictive efficiency were predominantly gained by using the amount of the number of failed tests. However, only slightly more variance could be explained. This confirms the results of former studies<sup>27</sup> and the appraisal that FCE information provides little additional information for predicting RTW status after rehabilitation. However, the results in this study provide new facts.

Because the cell's frequency distribution in the cross-tables was adequate only for persons employed at admission (n=123), the

sample was confined to those for estimating adjusted probabilities for a non-RTW isolated by subgroups. Patients with negative subjective prognosis of expected disability at work and high rates of preadmission sick leave had a high risk of non-RTW unrelated to the FCE information. In case all tests met the job requirements in this group, the predicted probability was .50 and thus higher than the sample's random probability of .38 (eg, 37.9% of the patients with non-RTW in the sample). For patients with a positive prognosis of expected disability at work and low rates of preadmission sick leave, the probability of non-RTW was lower than the sample's random probability of .38 unrelated to the FCE test results. Concerning these 2 groups, the FCE test did not provide any additional information predicting RTW.

However, for both groups with opposing information concerning the concurrent predictors (low rates of preadmission sick leave and negative prognosis of expected disability at work, as well as high rates of preadmission sick leave and positive prognosis of expected disability at work), a relationship between the number of failed tests and a probable non-RTW could be seen. For example, for 5 or more failed tests, the estimated probability was higher than the random probability for non-RTW.

Based on these results, a first proposal for a clinical prediction rule predicting RTW could be defined. As seen before, patients with consistent positive findings regarding the concurrent predictors (eg, low rates of preadmission sick leave and positive prognosis of expected disability at work) were categorized as "returners," whereas patients with consistent negative findings were "nonreturners." In case of opposing information, a limit of up to 5 failed tests could be used as a criterion for the definition of returners. More than 5 failed tests defined nonreturners. Using this prediction rule, a total of 76.9% (sensitivity: 69.7%, specificity: 80.0%) patients could be predicted correctly regarding RTW in the 1-year follow-up. Such a prediction rule revealed that especially for patients with oppos-

ing information concerning RTW prediction, FCE results could clarify a patient's situation. In contrast, FCE tests did not generate additional information when the prediction was distinctly positive or negative because of established predictors.

### Study Limitations

Some limitations of the study design should be considered. The sample size with valid cases was small, so that other potential predictors could not be sufficiently tested. Further, our results can only be compared limited to the results of others.<sup>25-27</sup> Our study referred to patients with MSDs (instead of low back pain only) and was focused on less complicated health problems. The patients reached an average of 4.1 points on the DOT scale, which equates to a heavy physical capacity. A similar study showed an average of 2.4 points only.<sup>20</sup> Also, only 4 failed tests (out of 25) confirmed that the patients in our sample had a heavy physical capacity. The study by Gross et al.<sup>24</sup> reported 8 and 9 deficits, respectively. Finally, the target variable's concept was to be reconsidered. "RTW in good health" was primarily important when considering the WHO definition for occupational participation. This definition implied an actual state (unemployment at follow-up) and an episode (at least 6 weeks of sick leave in the follow up-period) at the same time, which indeed improved the quality of the target variable.<sup>42</sup> To date, a combination of both findings was used in only some studies, so that it is uncertain whether the construct of occupational participation was met.

Consequently, in practice, it may be reasonable to use FCE results in addition to self-reported measures. It was shown that an FCE reduced to a few tests had the same predictive quality as the complex 2-day version.<sup>45</sup> So far, it has been debated whether the advantage justifies the high costs of FCE tests.<sup>28</sup> Considering the complexity of the prognosis of RTW, the use of FCE tests should be supported. A substantial improvement of prediction accuracy by using the amount of failed tests and a person's prognosis should be valued higher than the moderate quality of the model. Whether this will also lead to more effective and efficient rehabilitation programs is a subject for further (cost-effectiveness) research.

Because we used a sample with MSD without a specific disorder, it should be analyzed which other self-reported measures could be used in a clinical setting. Possibly there are more or fewer differences in specific disorders. Furthermore, it remained unclear to what extent the information of a later RTW is helpful in the following therapeutic process.

Finally, further research is needed for the use of FCE tests in other settings. It could be assumed that the effectiveness of FCE tests predicting RTW decreases by application to different populations—for example, to a population of white collar workers whose job requirements are different.

### CONCLUSIONS

This study indicated that IWS FCE could be used for patients whose occupational future was unclear, even if relevant information was applied. Information taken from IWS FCE tests was useful for the prediction of future occupational participation in a restricted manner. Because of economic parameters, alternative predictors should also be considered in practice. Nevertheless, relevant additional information could be gained. By combining these findings, the application of FCE tests could be carried out for patients with opposing occupational prognosis with the aim to improve the quality of the RTW prediction.

Finally, in the rehabilitation process, this could lead to an effective FCE-based prediction rule that could be used for the

identification of patients facing relevant problems in RTW despite conducting rehabilitation. In this context, FCE tests could play a major role in the inpatient and outpatient setting of rehabilitation more than they currently do.

### References

1. Maniadakis N, Gray A. The economic burden of back pain in the UK. *Pain* 2000;84:95-103.
2. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in the Netherlands. *Pain* 1995;62:233-40.
3. Frymoyer JW, Cats-Baril WL. An overview of the incidences and costs of low back pain. *Orthop Clin North Am* 1991;22:263-71.
4. Crook J, Moldofsky H, Shannon H. Determinants of disability after a work related musculetal injury. *J Rheumatol* 1998;25:1570-7.
5. Burdorf A, Naaktgeboren B, Post W. Prognostic factors for musculoskeletal sickness absence and return to work among welders and metal workers. *Occup Environ Med* 1998;55:490-5.
6. Mau M, Merkesdal S, Busche T, Bauer J. [Prognosis of labour force participation after multidisciplinary outpatient and inpatient rehabilitation for chronic back pain] [German]. *Rehabilitation (Stuttg)* 2002;41:160-6.
7. Turner JA, Franklin G, Fulton-Kehoe D, et al. Prediction of chronic disability in work-related musculoskeletal disorders: a prospective, population-based study. *BMC Musculoskelet Disord* 2004;5:14.
8. Schonstein E, Kenny DT, Keating J, Koes BW. Work conditioning, work hardening and functional restoration for workers with back and neck pain. *Cochrane Database Syst Rev* 2003;(1):CD001822.
9. Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. *Cochrane Database Syst Rev* 2002;(1):CD000963.
10. Harten JA. Functional capacity evaluation. *Occup Med* 1998;13:209-13.
11. Menard M, Gorman D. Work capacity evaluation. *N Z Med J* 2000;113:335-7.
12. Jette A. Functional capacity evaluation: an empirical approach. *Arch Phys Med Rehabil* 1980;61:85-9.
13. Smith S, Cunningham S, Weinberg R. The predictive validity of the functional capacities evaluation. *Am J Occup Ther* 1986;40:564-7.
14. Verband Deutscher Rentenversicherungsträger (VDR). FCE-Studie: FCE-Systeme zur Beurteilung der arbeitsbezogenen Leistungsfähigkeit. Bestandsaufnahme und Experteneinschätzung. DRV-Schriften: Frankfurt/Main; 2003.
15. Reneman M, Brouwer S, Meinema A, Dijkstra P, Geertzen J, Groothoff J. Test-retest reliability of the Isernhagen Work Systems Functional Capacity Evaluation in healthy adults. *J Occup Rehabil* 2004;14:295-305.
16. Reneman M, Dijkstra P, Westmaas M, Göken L. Test-retest reliability of lifting and carrying in a 2-day functional capacity evaluation. *J Occup Rehabil* 2002;12:269-75.
17. Brouwer S, Reneman M, Dijkstra P, Groothoff J, Schellekens J, Göken L. Test-retest reliability of the Isernhagen Work Systems Functional Capacity Evaluation in patients with chronic low back pain. *J Occup Rehabil* 2003;13:207-18.
18. Durand MJ, Loisel P, Poitras S, Mercier R, Stock SR, Lemaire J. The interrater reliability of a functional capacity evaluation: the physical work performance evaluation. *J Occup Rehabil* 2004;14:119-29.
19. Asante AK, Brintnell ES, Gross DP. Functional self-efficacy beliefs influence functional capacity evaluation. *J Occup Rehabil* 2007;17:73-82.

20. Reneman MF, Jorritsma W, Schellekens JM, Goeken LN. Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. *J Occup Rehabil* 2002;12:119-29.
21. Gross D, Battié M. Construct validity of a kinesiophysical functional capacity evaluation administered within a worker's compensation environment. *J Occup Rehabil* 2003;13:287-95.
22. Streibelt M, Hansmeier T, Müller-Fahrnow W. [Effects of work-related medical rehabilitation in patients with musculoskeletal disorders] [German]. *Rehabilitation (Stuttg)* 2006;45:161-71.
23. Gouttebarga V, Wind H, Kuijer P, Frings-Dresen M. Reliability and validity of Functional Capacity Evaluation methods: a systematic review with reference to Blankenship System, Ergos work simulator, Ergo-Kit and Isernhagen work system. *Int Arch Occup Environ Health* 2004;77:527-37.
24. Gross DP, Battie MC, Cassidy JD. The prognostic value of functional capacity evaluation in patients with chronic low back pain: part 1: timely return to work. *Spine* 2004;29:914-9.
25. Gross D, Battié M. Does functional capacity evaluation predict recovery in workers' compensation claimants with upper extremity disorders? *Occup Environ Med* 2006;63:404-10.
26. Gross D, Battié M. The prognostic value of functional capacity evaluation in patients with chronic low back pain: part 2: sustained recovery. *Spine* 2004;29:920-4.
27. Gross D, Battié M. Functional capacity evaluation performance does not predict sustained return to work in claimants with chronic back pain. *J Occup Rehabil* 2005;15:285-94.
28. Cutler RB, Fishbain DA, Steele-Rosomoff R, Rosomoff HL. Relationships between functional capacity measures and baseline psychological measures in chronic pain patients. *J Occup Rehabil* 2003;13:249-58.
29. Matheson LN, Isernhagen SJ, Hart DL. Relationships among lifting ability, grip force, and return to work. *Phys Ther* 2002;82:249-56.
30. Bürger W, Dietsche S, Morfeld M, Koch U. [Multiperspective estimates on the probability of patient return to work following orthopaedic rehabilitation: findings and predictive relevance] [German]. *Rehabilitation (Stuttg)* 2001;40:217-25.
31. Kuijer W, Groothoff J, Brouwer S, Geertzen J, Dijkstra P. Prediction of sickness absence in patients with chronic low back pain: a systematic review. *J Occup Rehabil* 2006;16:430-58.
32. Reneman M, Kool J, Oesch P, Geertzen J, Battié M, Gross D. Material handling performance of patients with chronic low back pain during functional capacity evaluation: a comparison between three countries. *Disabil Rehabil* 2006;28:1143-9.
33. Streibelt M. Aktivität und Teilhabe. Ein Beitrag zur Wirksamkeit berufsbezogener Maßnahmen in der medizinischen Rehabilitation der Rentenversicherung. Lengerich et al: Pabst Science Publishers; 2007.
34. Reneman MF, Kuijer W, Brouwer S. Item validity of the physical demands from the Dictionary of Occupational Titles for FCE of clients with chronic back pain. *Work* 2006;26:229; author reply 231-3.
35. Kersnovske S, Gibson L, Strong J. Item validity of the physical demands from the Dictionary of Occupational Titles for functional capacity evaluation of clients with chronic back pain. *Work* 2005;24:157-69.
36. Ware J, Sherbourne C. The MOS 36-item short-form health survey (SF-36), I: conceptual framework and item selection. *Med Care* 1992;30:473-83.
37. Tait RC, Chibnall JT, Krause S. The Pain Disability Index: psychometric properties. *Pain* 1990;40:171-82.
38. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. *Pain* 1999;83:157-62.
39. Reneman MF, Joling CI, Soer EL, Goeken LN. Functional capacity evaluation: ecological validity of three static endurance tests. *Work* 2001;16:227-34.
40. Steenstra IA, Verbeek J, Heymans M, Bongers P. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occup Environ Med* 2005;62:851-60.
41. World Health Organization. International classification of functioning, disability and health. Geneva: World Health Organization; 2001.
42. Dionne C, Bourbonnais R, Frémont P, Rossignol M, Stock S, Larocque I. A clinical return-to-work rule for patients with back pain. *CMAJ* 2005;172:1559-67.
43. Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.
44. Menard S. Applied logistic regression analysis. Series on quantitative applications in the social sciences, series no. 07-106, ed. Paper Sage University. Thousands Oaks: Sage; 1995.
45. Gross D, Battié M, Asante A. Evaluation of a short-form Functional Capacity Evaluation: less may be best. *J Occup Rehabil* 2007;17:422-35.